AN AUTOMATED SUBJECTIVE E-EVALUATION APPROACH USING ARTIFICIAL INTELLIGENCE

[1]R.Lakshmi,[2]K.Usharani,[3]P.Divya Bharathi,[4]R.Poornima @ Priyanka

[1]Professor,[2]Assistant Professor,[3]Assistant Professor, [4]Assistant Professor
[1]Department of Computer Science and Engineering,
[1]K.L.N. College of Engineering, Sivagangai, India

Abstract :   To ease the strain of the examination evaluation process, online tests and examinations are growing in popularity today. The questions on the online tests might be either objective or multiple-choice. Due to the difficulty and effectiveness of the review procedure, however, subjectively based questions and answers are not included. For universities and academic institutions, an automatic answer checker tool that verifies written responses and assigns grades in a manner comparable to that of a human person will be more beneficial. The current online exams are conducted and evaluated on machines which can contain only objective questions and there is no provision to extend these into subjective questions. To overcome this problem, a novel approach based on artificial intelligence (AI) is proposed in this paper to check subjective responses by awarding marks to the exam taker. The proposed system is based on keyword search algorithm that searches keyword provided by admin in the database and stemming algorithm that is used for linguistic normalization to evaluate. As a result of this artificial intelligence-based online answer evaluator, the evaluator's time and energy can be conserved with improved work efficiency.

Index Terms - **Evaluation Process, Online Tests, Artificial Intelligence, Subjective Responses.**

I. INTRODUCTION

In an online learning environment, it is crucial to evaluate the students' understanding of the subjects being taught. The system can modify or update the pedagogy based on the student's learning, suggest areas for additional research, and assess the students' performance. Many e-learning platforms use tests, quizzes, and assignments to evaluate students' learning. Based on how computers are utilized in the learning process, the educational system can be categorized into two categories: 1) traditional education (no computers used) and 2) e-learning, technology-enabled education (using information and communication technology). There will be direct communication between student and teacher under the conventional approach. The teacher observes and evaluates the student's understanding in addition to teaching. The education system underwent significant change over a long period of time as a result of developments in internet and communication technology. The extent to which computers are employed is the primary factor in evolution.

 One of the key benefits of e-learning is that it may provide adaptive learning, allowing instructors to change and distribute course contents based on how well students are currently doing. For a variety of reasons, regular knowledge assessments in e-learning must be conducted using various test types. Different assessments are used for the online test, however; this still is a thrust area for the researchers to propose novel E-assessment method. Due to the exam process is difficult and effective, subjective questions and answers have not been included in existing methods. It is directly affects the questions and answers related to the topic. Therefore, it is necessary to propose an E-assessment method based on artificial intelligence techniques to effectively verify subjective questions and answers. In this paper, a novel approach based on artificial intelligence (AI) is proposed to verify subjective responses by scoring the responses.

The proposed system is based on a keyword search algorithm that searches for keywords provided by the administrator in the database, and the stemming algorithm is used to standardize the language being evaluated. Through this AI-based online answer review tool, reviewers' time and energy can be saved through improved work efficiency.

II. RELATED WORK

Different researchers have different ways of defining e-assessment. Below are some definitions of e-assessment: according to Crisp (2011); Howarth (2015); and the Office for the Payment of Qualifications and Examinations (Ofqual), e-assessment involves the use of any technological device to create, deliver, store and/ or communicate assessment scores and student feedback. Examples of devices that can be used to create and perform electronic assessment tasks include laptops, desktops, smartphones, iPads, Android tablets, and more. Various multimedia formats such as Word documents, portable document formats (e.g. PDF), videos, images, simulations or games can be used to support e-assessment. With the advent of new technology, students' abilities and skills can be tested through electronic assessment (Crisp, 2011). JISC (2007), states that evaluation can be defined as "an end-

to-end electronic evaluation process" in which "information and communication technology (ICT) is used to present performance evaluation and record the responses". These processes involve the participation of all students, teachers, organizations and society. Ghanbarpour and H. Naderi (2010) , is proposed an attribute-specific ranking method based on language models to rank candidate answers according to their semantic information up to the attribute level.

III. E-ASSESSMENT

An assessment is sometime using the online assessment or e-assessment. Questionnaires and online tests can be done at home or in a controlled environment. Online assessments are similar to offline assessments but allow students to take the test or assessment through an internet-connected device. To take an online test or quiz, students must log into the assessment platform using their device (ideally a laptop or desktop). Once logged in, they can answer test questions - which can be multiple-choice, short answer, fill-in-the-blank, long answer, or a combination of these. Once the assessment is completed, the results are automatically saved and can be reviewed by teachers. Online assessments help measure a student's knowledge, understanding of the subject, and skills in the required area. Now that we know how online assessments work, let's look at some of the key differences between online and offline assessments.

The evaluation of examination is a process of assessing and measuring the educational system using many techniques like oral or viva-voce exam, subjective exam and objective exam etc. After conducting a survey on 500 people, we identified, objective exam is preferred to evaluate the students frequently due to its easy approach and less time consuming for the process of the evaluation of answer-sheet.

IV. PROPOSED METHOD

The current online exams are administered and evaluated on machines that can only include objective questions and there is no provision to extend it to include subjective questions. To overcome this problem, Artificial Intelligence (AI) based software application have been developed that automatically score users and verify subjective answers by matching answers in a database with answers written by test takers. An algorithm called "Stemming" is added. This identifies common terms and the semantic information present in the keywords. This extends the evaluation process to include awarding grades even when synonyms exist.

The PageRank algorithm has been added to help search engines find the best websites. Similarly, this algorithm evaluates student answers based on whether all keywords or only three to four keywords are present in the student's answer. The system evaluates responses to subjective questions by determining the match rate of keywords in instructor and student responses. Match rates are achieved based on semantic and document similarity. Exam scoring systems help instructors improve efficiency and productivity in administering and grading exams while saving time, cost, and resources.

4.1 Keyword Search Algorithm

Keyword search algorithm is a one of the procedure to find documents that have user's keywords. Normally, a search algorithm is using a full text that includes a list of user's search key words to identify the relevant documents that have the user's key words. If there is a keyword such as e-discovery, the search would be for that term and only that term . A straight keyword search methods only capture e-discovery, not electronic discovery, e-Discovery, or E discovery. Basic keyword search methods have some limitations as they may not discover word variants. Additionally, if there are typos or misspellings, such as abbreviations like "e-discovery" or "e-disco," these search terms will be missed by basic keyword search technology.

4.2 Stemming Algorithm

A stemming algorithm is a process of linguistic normalization, in which the variant forms of a word are reduced to a common form, for example,

connection
connections
connective → connect
connected
connecting

It is important to appreciate that we use stemming with the intention of improving the performance of IR systems.

4.3 PAGERANK Algorithm

In the search engine result, the Google Search is used the PageRank algorithm to rank the web pages. To compute the importance of the website. the PageRank algorithm works by counting the number of quality links to that page. The important websites are likely to receive more number of quality links from other websites. It is also used to rank keywords based on how important it is.

V. SYSTEM FRAMEWORK

5.1 Generation of the Question Base

The source of a question is either a Subject Expert or soft copies of previous papers, if available. Questions are preserved as cases within a question base. The type of the question component can vary depending on the answer expected. Specifically, these may be objective, short, descriptive, or short notes. Further, the distribution of marks and difficulty level need to be provided as well.

5.2 Generation of the Answer Component Base

Key points of answers to each question are actually housed within an Answer Component Base which is linked to the Answer Component List within the Question base by the index to a hash table representing the Answer Component Base.

5.3 Formation of Question Paper and Corresponding Temporary Paper Base

Papers are set automatically by choosing questions from the question base depending on selection criteria provided by the subject expert or any authorized person.

5.4 Guided Evaluation of Answer Scripts

This part which requires manual intervention of the examiner, who needs to assess the examinee's answers by consulting the answer components as preserved in the temporary paper base. Here, the examiner needs to check and resolve any semantic issues such as usage of synonyms and juxtaposition of key points. Marks are also allocated under system guidance as each answer component is associated with a percentage of the total marks of the question under consideration.

5.5 Generation of Score Sheets

A score file is automatically generated for each examination at the time of the guided evaluation process. Later, this may get updated during the reward and penalization processes. The evaluator can generate a final score sheet as and when required.

VI. SYSTEM DESCRIPTION

6.1 Pattern Generation Module

At first each question is converted into an affirmative sentence and places the template in place of the appropriate answer. For Example, Question: Where did "Brand XXIV" take place? Template: "Brand XXIV" took place in.

6.2 Textual Entailment Module

The various unigrams in the hypothesis for each Supporting Text (T) and Expected Hypothesis (H) pair are checked for their presence within the text. WordNet synsets for each unmatched unigrams in the hypothesis are found. If any one synset for the expected hypothesis (H) unigram matches with any synset of a word in the supporting text (T) then the hypothesis unigram is considered as a successful WordNet based unigram match.

6.3 Question-Answer Type Analysis Module

The questions are pre-processed by using Stanford Dependency parser. The question type and its expected answer type are generally identified by looking at the question keyword.

6.4 Answer Validation Module

If in the Expected hypothesis (H) NE is present then we check the result of the NER module. If NER module generates "VALIDATED" tag to the answer, then the results of the Textual Entailment module, Question-Answer Type Analysis module and Chunk & Dependency relations module are checked. If all these modules generate the "V ALIDA TED" result, the answer is tagged as "V ALIDA TED". Otherwise, the answer is tagged as "REJECTED".

VII. EXPERIMENTAL RESULTS

7.1 Preprocessing

The inputs to the module are the reference answers provided by the instructor and student answers. The module removes stop words, punctuations, and prepositions from answers. The output of this module is the cleaned answer. This process is done manually by using replace keyword in .net language. The preprocessed data can check the match easily reducing the complication of the further process. The frame work for preprocessing is given in Figure 7.1.

7.2 Keyword Expansion

The first task was to get synonyms in the wordnet, which is a lexical database of English available in the NLTK package (Natural language toolkit). Stemming algorithm is implied where all the similar words are made as one. For example, the term connective, connections, connected is all recognized and made as connect. This is done by removing quotations, special character and also spaces. It is given in the Figure 7.2.
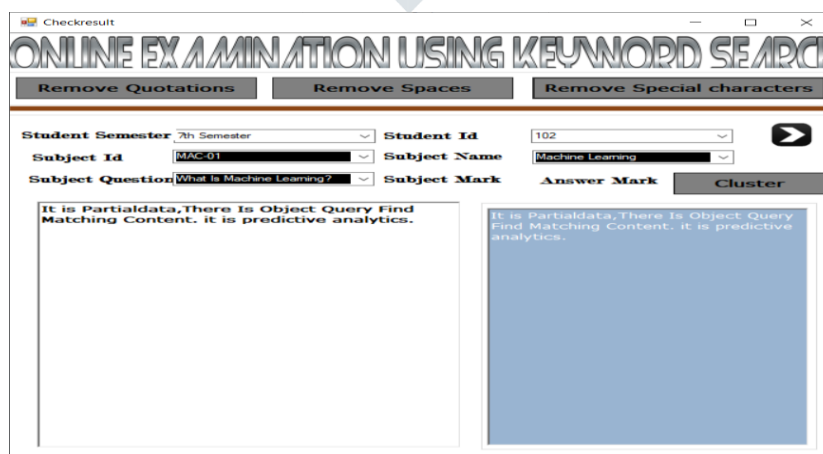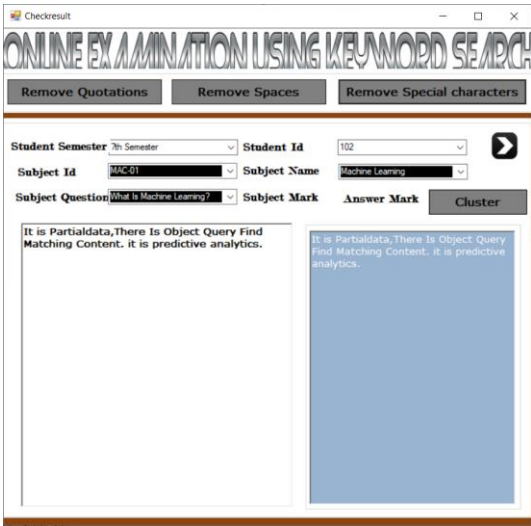

Figure 7.1: Preprocessing

Figure 7.2: Keyword Expansion

7.3 Matching Identifier

The module achieves two tasks: the first task is to convert the text to a vector and the second task is to compute cosine similarity between two vectors which is done using Keyword search algorithm. Then, the FRECCA (Fuzzy Relational Eigenvector Centrality-based Clustering) algorithm is applied for sentence level clustering which means the count of word and other is calculated after which the keyword search will check the match. The resultant frame work is shown in the Figure 7.3.
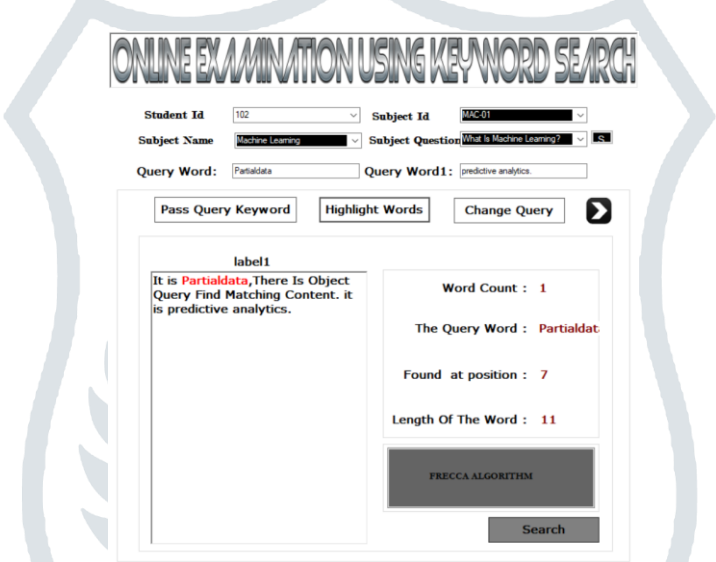

Figure 7.3: Matching Identifier

7.4 Grading Results

The module computes the full mark after assigning marks to each question. This is where the ranking algorithm will be implemented where based on the 2 keyword; the marks for 2m will be allotted. PageRank algorithm is used to allot marks accordingly. This frame work is shown in Figure 7.4.
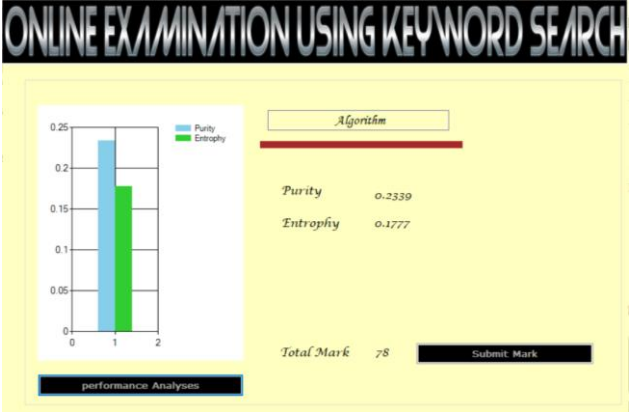

Figure 7.4: Grading Result

7.5 Evaluation

The module represents the result in graphical representation - the execution times, number of terms, frequency and precision in both student and staff login. For example, how many students have passed or failed will be displayed. The main process of the module is displaying the results in graphical representation. The Frame work of evaluation is shown in the Figure 7.5.
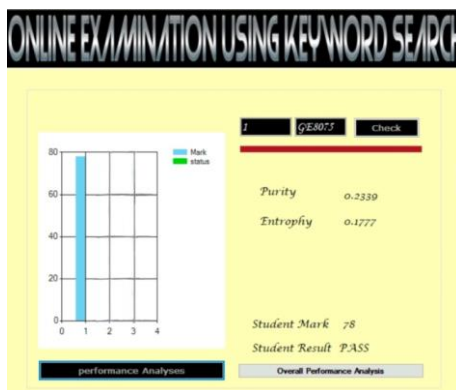

Figure 7.5: Evaluation

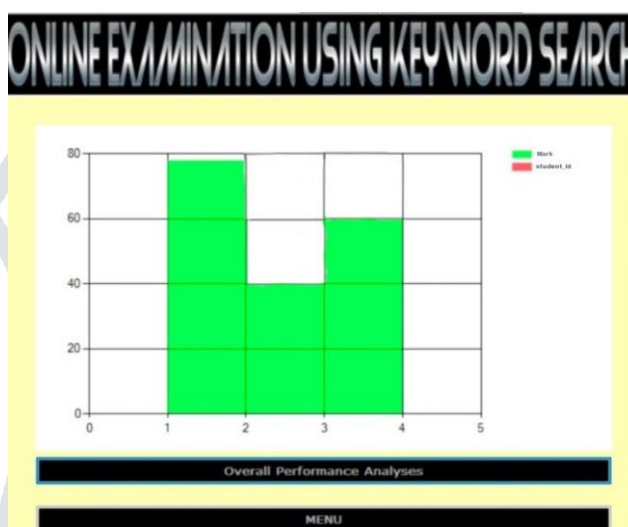The overall performance analysis is given in the Figure 7.6.

Figure 7.6: Overall Performance Analysis

VIII. CONCLUSION

The answer sheet is widely used for student performance in exam in school and college. The main approach is to evaluation is efficient and reliable. An automatic answer sheet checker verifies the answer sheet and written mark as similar to human being. This software is built to check the subjective answer. Examinations play a very important role in colleges, universities and various other educational institutes. Many educational institutes have their examinations conducted online, but these exams only contain multiple choice questions which are providing to be very efficient in testing the student's aptitude, on the other hand fail to measure the conceptual knowledge a student or learner must possess. Therefore, subjective answer must be included in online examinations.

The proposed system evaluates the answer based on the keywords. By comparing the standard answer and the student's answer marks is obtained if the student utilizes all the keywords mentioned in the standard answer. Hence the said system could be of great utility to the educators whenever they need to take a quick test for revision purpose, as it saves them the trouble of evaluating the bundle of papers.

IX. FUTURE ENHANCEMENT

The system can be implemented to 10-mark questions also. As of now only 2 marks is evaluated based on 2 keywords. One short fall of the present system is that till now it can only evaluate 2 marks. In order to be a full-fledged AI tool, it needs to allow images and symbols at every phase. Another limitation of the system is its inability to assess semantic aberrations appearing within the student's response. Scope of automation in the checking and marking procedure may also be explored further. All these shortcomings need to be taken care of by employing state-of-the-art multimedia mining, Natural Language Processing and Computer-Vision techniques involving image processing as well as optical character recognition, in future models.

REFERENCES

[1]. Agrawal S., Chaudhuri S., and Das G. 2002. DBXplorer: A System for Keyword-Based Search over Relational Databases. Proc. of the 18th International Conference on Data Engineering, IEEE Computer Society, pp. 5-16.

[2]. Ani Thomas, MKKowar and Sanjay Sharma. 2008. Intelligent Fuzzy Decision Making For Subjective Answer Evaluation using Utility. Emerging Trends in Engineering and Technology 2008 ICETET '08 First International conference.

[3]. Coffman J. and Weaver A.C. 2014. An Empirical Performance Evaluation of Relational Keyword Search Techniques. IEEE Tran. on Knowledge and Data Engineering, vol. 26, no. 1, pp. 30-42.

[4]. Dhokrat A., Hanumant G., and Namrata Mahender R.C. 2012. Automated Answering for Subjective Examination", International Journal of Computer Applications, Volume 56, No.14, pp: 14-17.

[5]. Ghanbarpour A and Naderi H. 2019. A Model-based Keyword Search Approach for Detecting Top-k Effective Answers. The Computer Journal, vol. 62, no. 3, pp. 377-393.

[6]. Ghanbarpour A and Naderi H. 2020. An Attribute-Specific Ranking Method Based on Language Models for Keyword Search over Graphs. IEEE Transactions on Knowledge and Data Engineering, vol. 32, no. 1, pp. 12-25.

[7]. Gunjal,Mrunal A., Sayli M., Pawar M and Prakash J.Kulkarni. 2015. Evaluation of Subjective answers using GLSA enhanced with contextual synonyms. International Journal on Natural Language processing Computing(INLC) Vol 4. No1.

[8]. Hristidis V. and Papakonstantinou Y. 2002. Discover: keyword search in rela-tional databases. Proc. of the 28th international Conference on Very Large Data Bases, VLDB Endowment, Hong Kong, China, pp. 670-681.

[9]. Lakshmi R and Baskar S. 2019. DIC-DOC-K-means: Dissimilarity-based Initial Centroid selection for DOCument clustering using K-means for improving the effectiveness of text document clustering", Journal of Information Science, vol. 45, Issue 6, pp.818-832.

[10].Lakshmi R and Baskar S. 2021. Efficient Text Document Clustering With New Similarity Measures. International Journal of Business Intelligence and Data Mining, vol. 18, no, 1, pp.49-72.

[11].Park C.S. and Lim S. 2015. Efficient processing of keyword queries over graph databases for finding effective answers Information Processing & Management, vol. 51, no. 1, pp. 42-57.

[12].Parkavi A., Bhavana G Dongre., Gayathri Devi., Spoorthy S.Hathwar., and PatnamPranitha. 2013. Algorithm For Automatic Evaluation Of Single Sentence Descriptive Answer. International Journal Of Inventive Engineering and Sciences (IJIES) ISSN:2319- 9598, Volume-1, Issue-9.

[13].Roy C and Chaudhuri C. 2018. Case Based Modeling of Answer Points to Expedite Semi-Automated Evaluation of Subjective Papers. 2018. IEEE 8th International Advance Computing Conference (IACC), Greater Noida, India, pp. 85-90.

[14].Sheeba Praveen. 2014. An Approach to Evaluate subjective Questions for Online Examination System. International Journal Of Innovative research in Computers and communication Engineering Volume-2, Issue 11.